



THE BATTLE OF NEIGHBOURHOODS

IBM PROFESSIONAL CERTIFICATE IN DATA SCIENCE

BUSINESS PROBLEM

- The main objective of this Project is to analyse Madrid city and to find the best position for opening a new restaurant.
- In the study it will be analysed both position and type of restaurant.
- **IMPORTANCE:** Choosing the right location and specialization for a restaurant can make the difference between a huge success or a big failure in the business

MADRID CITY

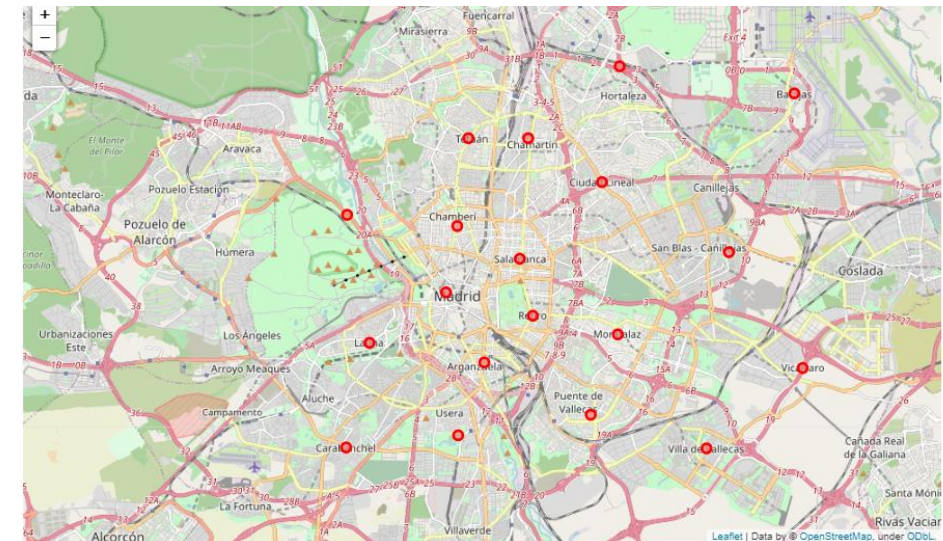
Madrid is the capital of Spain. The city of Madrid has 3.2 million of inhabitants that are distributed in 21 boroughs. The city has a large amount of restaurants, including different types of food such as typical spanish food, asian food, mexican food, etc



DATA ACQUISITION

- Wikipedia -> Boroughs of Madrid information: https://es.wikipedia.org/wiki/Anexo:Distritos_de_Madrid
- Foursquare API: Information of venues in the are of the Madrid boroughs.

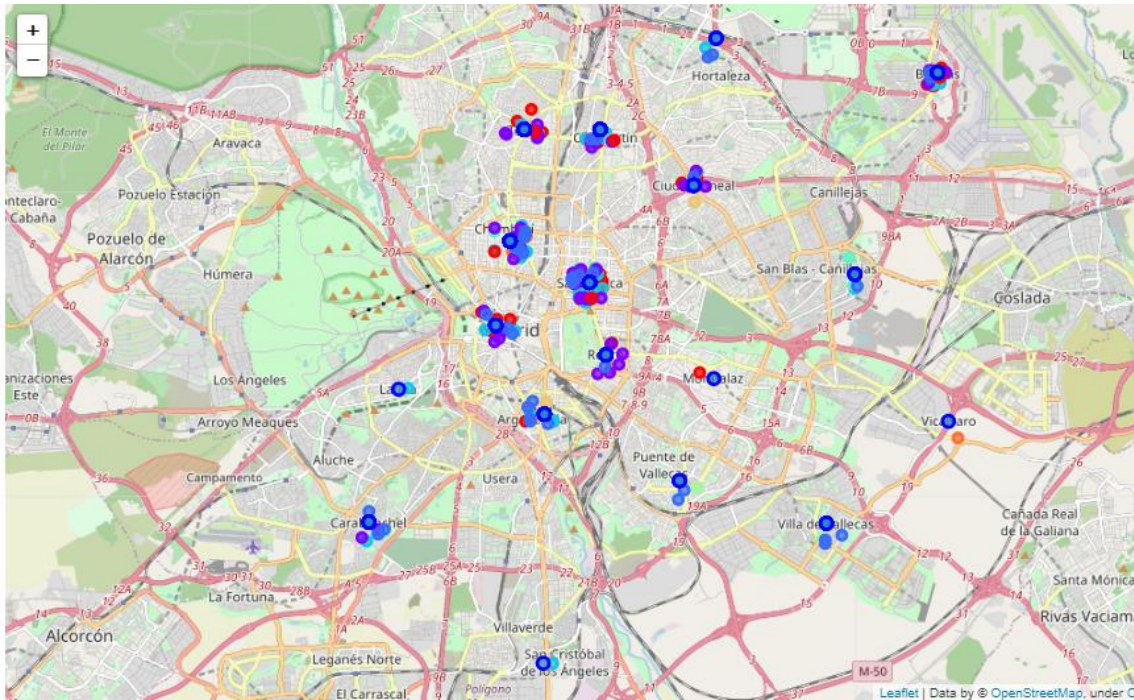
	Borough	Area[ha]	Poblation	Density[pob/ha]	Latitude	Longitude
0	Centro	522.82	131928	252.34	40.417653	-3.707914
1	Arganzuela	646.22	151965	235.16	40.398068	-3.693734
2	Retiro	546.62	118516	216.82	40.411150	-3.676057
3	Salamanca	539.24	143800	266.67	40.427045	-3.680602
4	Chamartin	917.55	143424	156.31	40.460764	-3.677534
5	Tetuan	537.47	153789	286.13	40.460821	-3.699520
6	Chamberi	467.92	13401	293.64	40.436247	-3.703830
7	Fuencarral-El Pardo	23783.84	238756	10.04	40.556346	-3.778591
8	Moncloa-Aravaca	4653.11	116903	25.12	40.439495	-3.744204
9	Latina	2542.72	233808	91.95	40.403532	-3.736152
10	Carabanchel	1404.83	243998	173.68	40.374211	-3.744676
11	Usera	777.77	134791	173.30	40.377602	-3.703378
12	Puente de Vallecas	1496.86	227595	152.05	40.383553	-3.654535
13	Moratalaz	610.32	94197	154.34	40.405933	-3.644874
14	Ciudad Lineal	1142.57	212529	186.01	40.448431	-3.650495
15	Hortaleza	2741.98	180462	65.81	40.480825	-3.644027
16	Villaverde	2018.76	142608	70.64	40.343092	-3.694119
17	Villa de Vallecas	5146.72	104421	19.86	40.373958	-3.612163
18	Vicalvaro	3526.67	70051	19.86	40.396584	-3.576622
19	San Blas-Canillejas	2229.24	154357	69.24	40.428919	-3.604002
20	Barajas	4192.28	46876	11.18	40.473318	-3.579845



DATA SCIENCE METHODOLOGY

- **Scrapping data** from Wikipedia webpage, obtaining the boroughs with their área and population information
- Preprocess the data to remove not desirable variables
- Obtain the location of the boroughs using the **geopy library**
- Use the **Foursquare API** to find the restaurants in the boroughs areas
- Classification of the restaurants using the **K-Means algorithm**
- Plot the restaurant clusters in the Madrid map using the **Folium** library

RESULTS



Clusters of restaurants located on each borough

Cluster	Restaurant Type
0 (Red)	Italian-Argentinian-Indian-American Restaurants
1 (Purple)	Spanish Restaurants
2 (Darkblue)	Tapas Restaurants
3 (Clearblue)	Generic Restaurants
4 (Cyan)	Seafood Restaurants
5 (Green)	Asian Restaurants
6 (Yellow)	Mexican Restaurants
7 (Orange)	Mediterranean Restaurant

CONCLUSIONS AND OBSERVATIONS

1. The Centro borough which is the most touristic concentrate a high density of Spanish and Tapas Restaurant where typical Spanish meals are provided.
2. In general, taking in account all the boroughs, the blue-purple colour is the predominant, which means that typical Spanish restaurant are the most extended.
3. If we continue focusing in the most central boroughs the most interesting restaurant possibilities are international restaurants such as Mexican or Italian due the lower amount of establishment focused on that foods.
4. In areas like Chamartin or Tetuan it can be observed a lack of Asian restaurant which means that opening one of this establishments can be good decision based also in the population of these areas.
5. Due to the high amount of borough, quite considerations can be extracted attending to the classification done, with just looking the plotted map.